DOCUMENT RESUME

ED 462 442		TM 033 706

AUTHOR	Capraro, Robert M.; Graham, James M.
TITLE	A Tutorial on How Reliability Is Estimated within Structural Equation Models.
PUB DATE	2002-02-10
NOTE	48p.; Paper presented at the Annual Meeting of the Southwest Educational Research Association (Austin, TX, February 14-16, 2002).
PUB TYPE	Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Error of Measurement; *Estimation (Mathematics); Reliability; *Scores; *Structural Equation Models

ABSTRACT

This paper illustrates first how estimated Structural Equation Modeling (SEM) measurement error variances are actually estimates of score reliabilities. The major advantage of SEM over other analytic methods is that it accounts for measurement error. Score reliabilities are estimated as part of structural modeling, so that structural models test both substantive hypothesis and measurement models. The paper also demonstrates how changing the estimated reliabilities of the SEM observed variables not only changes estimated reliabilities, but can also change all the other parameters throughout the analysis. A heuristic data set is used to illustrate these changes. One appendix contains the Statistical Package for the Social Sciences syntax used for the illustration, and the other shows data from a previous study used for illustrative purposes. (Contains 4 tables, 6 figures, and 33 references.) (Author/SLD)

Running head: Reliability Estimation in SEM

A Tutorial on How Reliability is Estimated Within Structural Equation Models

Robert M. Capraro James M. Graham

Texas A&M University

Paper presented at the Southwest Educational Research Association, February 14-16, 2002, Austin, TX. *Inquires concerning this paper should be addressed to the author at College of Education, Texas A&M University, 4232 TAMU, College Station, TX, 77843-4232, (979) 845-8007. E-mail: rcapraro@coe.tamu.edu*

Abstract

The paper illustrates first how estimated-Structural Equation Modeling (SEM) measurement error variances are actually estimates of score reliabilities. It also demonstrates how changing the estimated reliabilities of the SEM observed variables not only change estimated reliabilities, but also can change all the other parameters throughout the analysis.

A Tutorial on How Reliability is Estimated Within Structural Equation Modeling

As Thompson (1991) noted, "Multivariate methods best honor the reality to which the researcher is purportedly trying to generalize." (p. 80) Canonical correlation analysis (CCA) has been found to subsume all other parametric analyses such as *t*-tests, ANOVA, *regression, and* X^2 (Frederick, 1999; Knapp, 1978). In seminal work Jöreskog and Sörbom (1979) indicated that SEM is an even more general case of the GLM, subsuming all other cases of the as special cases (Capraro & Capraro, 2001; Knapp, 1978; Thompson, 1991).

Because SEM subsumes all other parametric statistical analyses it provides some interesting options for the researcher. First, all other analyses (e.g., *t*-tests, ANOVA, regression, MANOVA, and CCA) can be conducted as special cases in SEM. This is of conceptual interest but is often not a practical shortcut or an elegant solution that efficiently answers the research question at hand. When the research questions deal with understanding the underlying structure (EFA) of a set of items or in confirming a theory from a set of data (CFA) SEM can provide truly elegant and unique solutions (Stevens, 1996).

Structural Equation Modeling

The work of Jöreskog in the early 70's led to statistical theory and algorithms used in the analysis of linear structural relation models known today as structural equation modeling. In SEM, the models include unobservable variables identified as latent constructs, defined by observed variables, which by theory fit the construct(s). Measurement error, reflecting score reliability is also typically estimated as a unique and an essential part of SEM analyses (Jöreskog, 1969, 1970, 1973, 1977; Jöreskog, & Goldberger, 1975; Jöreskog, & Sörbom, 1978).

Covariance structure modeling (CSM) or SEM is being used increasingly within the social sciences (Stevens, 1996). Indeed, it would be difficult to locate recent issues of social science journals in which some SEM applications were not reported. And one new journal--*Structural Equation Modeling: A Multidisciplinary Journal*--has been created that is devoted exclusively to SEM reports and issues. SEM has been termed "the single most important contribution of statistics to the social and behavioral sciences during the past twenty years" (Lomax, 1989, p. 171). Similarly, Stevens (1996) argued that SEM is ". . . one of the most important advances in quantitative methodology in many years" (p. 415).

Various authors have explained SEM in an accessible manner (cf. Thompson, 2000). It is important to recognize that SEM is the most general case of the General Linear Model (GLM), which means that SEM subsumes other methods (e.g., _t_ tests, ANOVA, regression, MANOVA, descriptive discriminate analysis, canonical correlation analysis) as special cases (Bagozzi, Fornell, & Larcker, 1981; Cohen, 1968; Knapp, 1978; Thompson, 1991).

SEM's major advantage over other analytic methods is that it accounts for measurement error. A major distinguishing feature of SEM is that score reliabilities are estimated as part of structural modeling. Thus, structural models test _both_ substantive hypothesis and measurement models. This may not be obvious to many applied researchers, because rather than estimating reliabilities directly, SEM estimates error variances instead.

The paper will illustrate first how these error variances are actually estimates of score reliabilities (Thompson, 2000). It will also be demonstrated how changing the estimated reliabilities of the SEM observed variables not only changes all the other parameters throughout the analysis. Because reliability estimation is such an important part of SEM, it is important for researchers to

understand how these estimates are incorporated into the models.

Reliability

Reliability can be viewed from several perspectives. Reliability can be considered as the portion of any variable that cannot be attributed to measurement error. At the level of scores of given individuals, reliability of the observed scores ("O" or "X") can be expressed as a finction of he each person's "true" or reliable score component ("T") and a corresponding measurement error score component ("E"), subject to the restriction that: $O_i = T_i + E_i$. For a set of scores, the reliability of the data (remembering that tests are <u>not</u> reliable, data are reliable) can be estimated by correlating the O_i and the T_i scores. Alternatively, score reliability can be estimated as $\frac{Var_{ti}}{Var_{oi}}$. The concept is simple but in the application one first notices the complexities for estimating the reliability. For a more rigorous treatment on reliability refer to Cronbach (1951), Arnold (1996), Reinhardt (2000), and Henson (2001).

As Thompson (1994) noted, "The same measure, when administered to more heterogeneous or more homogeneous sets of subjects, will yield scores with differing reliability"

(p. 839). Because of this phenomenon, researchers should always examine and report the reliability of their data in hand even for substantive studies. As the APA Task Force on Statistical Inference emphasized

> . . . it is important to remember that a test is not reliable or unreliable. Reliability is a property of the scores on a test for a particular population of examinees . . . Authors should provide reliability coefficients of the scores for the data being analyzed even when the focus of their research is not psychometric. (Wilkinson & APA Task Force on Statistical Inference, 1999, p. 596)

Similarly, Gronlund and Linn (1990) indicated that reliability is based on the results obtained from an evaluation instrument and not a property imbued at creation within the instrument itself. From this perspective, it is most appropriate to speak of reliability as a factor of test scores or of measurement rather than of the test or of the instrument.

Although many researchers refer to "the reliability of the test" in some very eloquent prose, this tendency may lead to the misconception that reliability is a property of instruments rather than a property of scores. This often naïve conception can result in researchers failing to

examine score reliability for their data (Capraro, Capraro, & Henson, 2001).

Carmines and Zeller (1979) referred to reliability as the part remaining after partitioning purely random error. Jöreskog and Sörbom (1993) and Hancock (1997) indicated that in SEM reliability is defined as the variance in the variable is not accounted for by measurement error. Raines-Eudy (2000) noted, "It [reliability] is commonly represented by the squared multiple correlation coefficient, which ranges from 0 to 1" (p. 128). Lomax (1989) "error variance is random, unsystematic and due to unreliability. . . Specific variance is non-random, systematic, reliable, and due to the particular selection of variables by the researchers" (p. 193).

Reliability Unavoided

Although many researchers choose not to report reliability coefficients for data in hand (Capraro, Capraro, & Henson, 2001), performing sophisticated statistical analyses such as SEM invokes the reliabilities of data in hand in the process. In the case of structural equation modeling (SEM), error variances are estimated and low error variances as compared to high error variances do produce very difference results.

Lomax (1989) found that

> . . . when the level of a specific variance (due to reliability) was systematically reduced for a single exogenous variable (but not for the other two), (1) error variance for that variable increased systematically since specific and error variance are analytically, inversely related . . . Thus, in those situations where indicator variable reliability differs substantially from unity, say 0.8 or less consideration needs to be given to taking the unreliability into account in the analysis; otherwise, certain important parameter estimates may be biased leading to possible model misrepresentation. (p. 193)

In classical statistical analyses such as ANOVA, multiple linear regression, and MANOVA do not directly consider reliability as part of the analysis. However, measurement error (reliability) in the classical statistical analyses does influences parameter estimates and effect size (Thompson, 1998). Selecting SEM as the analytic method, on the other hand, forces the researcher to confront the reality of reliability of the data in hand. Given the increased controversy over statistical significance testing and effect sizes (Thompson, 1996), it is important to understand that given a large enough sample

one would always achieve statistical significance whereas "effect-size measures do not rely on sample size, their use would be most beneficial for comparing findings from studies involving different sample sizes" (Devaney, 2001, p. 311). However, effect sizes are attenuated by reliability coefficients (Thompson, 1998). The effect size yield cannot exceed the limits imposed by the reliability coefficient.

Measurement Error in GLM and SEM

As previously discussed, reliability can be conceptualized as the proportion of "true", or non-error, variance in a given measured variable. Thus, the variance of a measured variable X is made up of both true variance and error variance, or as expressed in classical test theory, X = true + error. In graphical representations of SEM models, measured variables are typically shown as squares, while latent or synthetic variables are shown as circles. Direct relationships such as those between an independent and dependent variable in regression are shown as a single-ended arrow, while correlations are shown as a double-ended arrow. The classical test theory formula, re-expressed in terms of structural equation modeling, is shown in Figure 1. Figure 1 shows a measured variable 'X',

which is made up of two synthetic variables, the "true" variance and the error variance. The reliability of *X* could be expressed as the proportion of the variance of *X* in the "true" latent variable.

INSERT FIGURE 1 ABOUT HERE

The concept of dividing the variance of a measured variable into "true" and error synthetic variables pervades all general linear model (GLM) procedures. However, this process is only ever invoked for the dependant variable (or in the case of multivariate techniques, dependent variables), and the partitioning of the variance of the dependant variable is dictated by the relationship of the dependent variable with the independent variables. For example, the variance of the dependent variable in an ANOVA analysis is partitioned into the variance of the dependent variable explained by the independent variables (referred to as explained or between variance) and the variance of the dependent variable unexplained by the independent variables (referred to as unexplained, within, or error variance). The variance explained is comparable with the "true" variance discussed above; dividing the variance explained by thee total variance of the dependent variable results in the ANOVA eta-squared, an r-squared type effect size.

This same process holds true for all GLM procedures. In multiple regression, the dependent variable is divided into error and "true" variance. The non-error variance in regression is the variance predicted by the independent variables, sometimes referred to as the variance of $\hat{y}$ or y-hat. As with ANOVA, dividing the variance of y-hat by the total variance of the dependent variable results in the R^2 effect size of the regression model. In a bivariate correlation, the proportion of covariance (the variance shared by two variables) to the total variance results in a squared correlation; taking the square-root of this number gives you the correlation coefficient.

Similar latent variables are invoked in both GLM procedures and reliability processes. In both procedures, the proportion of true variance to total variance results in a r-squared statistic which is the focus of the analysis. In GLM procedures, this r-squared statistic is the effect size (e.g., eta-squared, R^2, etc.), and in a reliability analysis this r-squared statistic is the reliability coefficient (e.g., Cronbach's alpha). However, there are several important distinctions between these procedures. In GLM procedures, the partitioning of the dependent variable is dependent on the dependent variable's relationship with the independent variables. In

reliability estimation, the variables are partitioned as a function of internal consistency and measurement error. Additionally, GLM procedures only partition the variance of the dependent variable; all of the variance of the independent variables is implicitly assumed to be error-free.

The strength of SEM procedures lies in the ability to estimate the error variance in all variables included in the analysis, not just in the dependent variable as with other GLM procedures. Thus, it is possible to set the error variance for a given variable to be equal to the measurement error, or un-reliability, of that variable.

Heuristic Examples

A SEM model for conducting a multiple regression type analysis in AMOS 4.0 was used to for heuristic demonstration with the results contrasted against the results from a multiple regression analysis conducted in SPSS. Because SEM subsumes multiple regression it is possible to graphically demonstrate the differences in the results from a traditional multiple regression analysis that does not account for reliability and a SEM model that does account for reliability. Initially, a regression is

conducted using Statistical *Packages for the Social Sciences* v. 10 (SPSS) to demonstrate that SEM produces equivalent results. Second, two models are developed to illustrate the differences in the estimated parameters due to changes in reliability. For accessibility the Holzinger and Swineford (1939) data set was used so interested individuals can replicate the procedures contained herein and achieve the same results. Originally the data set contained 301 cases with 26 variables, but for illustration of the methods a subsample of the data set was used by selecting only those students enrolled in Grant-White school (n=145). The SPSS syntax is contained in Appendix A and the data are contained in Appendix B. The variables t5 (general information verbal test), t6 (paragraph comprehension test), t8 (word classification) and t15(memory of total numbers) were used to predict t9 (word meaning).

In the following example, data from the general information (GI), paragraph comprehension (PC), word meaning (WM), word classification (WC) and memory of total numbers (NR) were used (See Appendix B).

Regression

A regression analysis was conducted in SPSS using GI, PC, and NR, to predict WM; the results are contained in Table 1. The SPSS syntax used to run this analysis is presented in Appendix A. As shown in Table 1, the R^2 for this analysis is .635, with both GI and PC having standardized beta-weights greater than .4, and variable NR having a near-zero beta-weight. As previously discussed, the R-squared statistic represents the proportion of the variance in the dependent variable predicted by the independent variables. Thus, the independent variables account for 63.5% of the variance in the dependent variable WM.

INSERT TABLE 1 ABOUT HERE

The same regression model, expressed in graphical SEM terms is shown in Figure 2. In this model, the four measured variables are represented by squares. The dependent variable WM is shown partitioned into two latent variables, y-hat and error; note that this is identical to the classical test theory equation shown in Figure 1. The three independent variables are shown as correlated (indicated by the double-ended arrows) and as predicting y-hat, or the non-error portion of the dependent variable. In this model, the independent variables are not

partitioned into error and non-error terms—this is because multiple regression does not take into account error for the independent variables.

INSERT FIGURE 2 ABOUT HERE

The model shown in Figure 2 was then used as an input model in AMOS 4.0, a commonly used graphical SEM modeling program. Before the model was analyzed, a number of steps were taken to format the analysis. From the "View/Set" - "Analysis Properties" menu, the "Output formatting" tab was selected. On this screen, the boxes "standardized output" (instructing AMOS to provide both standardized and unstandardized estimates) and "squared-multiple-correlations" (instructing AMOS to report the proportion of non-error variance to the total variance for each measured variable partitioned into error and non-error terms) were checked. From the "Bias" tab, the use of unbiased covariance matrices for both input and analysis was specified. Finally, from the "estimation method" tab, the use of an unweighted least-squares method was specified.

The resulting standardized AMOS output from the input model shown in Figure 2 is shown in Figure 3. In Figure 3, the estimated correlations are shown printed next to double-ended arrows, standardized regression weights are

shown next to single-ended arrows, and the multiple-squared-correlations are shown next to any measured variable partitioned into error and non-error terms. The SEM results shown in Figure 3 are identical to the SPSS regression output shown in Table 1 and 2. The squared-multiple-correlation (or the proportion of non-error to total variance) of WM is shown in the AMOS output as .63, which is equal to the R^2 of .635 shown in the SPSS output, within rounding error. Because the non-error synthetic variable, y-hat, is made explicit in this SEM model, it is necessary to multiply the regression weight between an independent variable and y-hat by the regression weight between y-hat and the dependent variable to obtain the equivalent beta-weight for that independent variable. For example, the line between GI and y-hat is shown as having a weight of .56 and the line between y-hat and WM is shown as having a weight of .80. When these numbers are multiplied, the result is the beta-weight (.56 x .80 = .448) shown in Table 1, within rounding error.

As an aside, if "all implied moments" are requested from the Analysis properties/output menu, the AMOS text output will include the correlations between the dependant variables and y-hat, also known as structure coefficients. Structure coefficients, used in conjunction with beta-

weights, are an important part of interpreting regression results.

INSERT FIGURE 3 ANS TABLE 2 ABOUT HERE

Regression with Measurement Error

Next, a SEM model that accounts for measurement error was generated. This model accounts for measurement error of the measured variables; note the associated error variances in Figure 3. Table 3 contains a correlation matrix which confirms the results achieved from the SEM model.

While the regression model discussed above does not include error terms for the independent variables, the use of SEM procedures allows error terms to be estimated for all variables included in an analysis. When error terms for the independent variables in the regression model are added, the results are shown in Figure 4. In this model, the partitioning of WM into y-hat and error remains identical to the regression model shown above. However, instead of the independent variables being used to directly predict y-hat, each independent variable is given its own error term. All independent variables are shown as sharing a common "true", non-error term, and it is this non-error portion that is shown as being used to predict y-hat.

INSERT FIGURE 4 ABOUT HERE

It is important to note that the model shown in Figure 4 is <u>not</u> simply regression with error terms added to the independent variables. In fact, it is no longer regression at all, as regression does <u>not</u> allow error to be estimated for independent variables. By adding error terms to GI, PC, and NR, they cease to function as independent variables in a mathematics sense. That is, while the partitioning of the variable WM is dependant on its relationship with the other variables, the partitioning of each of the once independent variables is likewise dependant on that variable's relationship with all other variables in the analysis. Rather than regression, the model shown in Figure 4 is a confirmatory factor analysis model; an equivalent representation of this model could be obtained by removing the y-hat synthetic variable and creating a single-ended arrow pointing from the "true" synthetic variable to the WM measured variable.

Structural Equation Modeling

Because the addition of error terms to the independent variables in the regression model cause it to cease to function as a regression model, the SEM model shown in Figure 5 will be used for all of the following models. In

this model, the variables GI, PC, and NR are used to construct a latent variable, Var1. The variables WM and WC are used to construct a second latent variable, Var2. The relationship between the two latent variables, Var1 and Var2, is similar to the relationship between the independent and dependant variables in a regression model, as shown by the single-ended arrow between them. The following analyses use the same settings for AMOS described above, save that the covariance matrix to be analyzed and estimation discrepancy are set to maximum likelihood, a commonly used estimation method in SEM procedures.

A series of analyses using this input model will now be presented to show the impact that changing measurement error can have on statistical results. AMOS output diagrams are not presented for all models. For the sake of brevity, the AMOS output diagram for model 1 is shown, while the squared multiple correlations and fit indices for all other models are presented in tabular form.

INSERT FIGURE 5 ABOUT HERE

Model 1. SEM software programs such as AMOS allow the user to choose between fixing the variance of exogenous variables and regression weights to a value chosen by the user or estimating them as part of the analysis. By

double-clicking on a path or variable in AMOS, the user is given the option of setting the regression weight of a line or the variance of a latent variable to a value. This is done by assigning variance to the error term. The error term can be adjusted to demonstrate the effects of measurement error on the output results. A low variance reflects little or no measurement error in a given variable. A large variance set for the error term indicates that most or all of the variance in the measured variable is due to measurement error or score unreliability. The AMOS output shown in Figure 6 has allowed the variances of all error terms freed; that is, no regression weight or latent variable variances were fixed to a given number. The squared multiple correlations and several fit indices, the chi-square statistic, the root mean square error of approximation (RMSEA) and adjusted goodness of fit index (AGFI), for this model are summarized as model 1 in Table 2. In SEM, allowing all parameters to be estimated, as done in model 1, always results in the best fit possible for the data. The fit indices for this model therefore represent the best possible fit of the data to this model.

INSERT FIGURE 6 AND TABLE 2 ABOUT HERE

Model 2. Because SEM allows the user to specify the amount of error variance for a variable, it is possible to set the error terms for a given variable to be equal to the un-reliable portion of the measured variable variance, if the reliability coefficients are known. All of the variables from the Holzinger data-set are composite scores made up of multiple test items. The reliability coefficients for these variables have been published in several manuscripts. The reliability coefficients shown in Table 3 are taken from Gorsuch (1983). Because the reliability coefficient is the proportion of the total variance that is not due to measurement error, it is possible to determine the amount of error and non-error variances if the total variance is known. Table 4 shows the total variances of the sample data with the non-error variances (obtained by multiplying the total variance by the reliability coefficient) and error variances (obtained by multiplying the total variances by 1 - the reliability coefficient).

INSERT TABLE 3 ABOUT HERE

Model 2 was created by using the same input model shown in Figure 5, with the exception that the error

variances of all measured variables were set to be equal to the error variances determined by the reliability coefficients shown in Table 4. The results of this analysis are shown in Table 3. As seen here, the model has a considerably worse fit when compared to the model where error terms were left free to be estimated, model 1. Additionally, the squared multiple correlations have changed considerably for many of the variables.

Just as the variance for error can be set very low as in the previous example it can also be set to almost all the variance. This would result in most of the prediction power of the measured variable being consumed by error with little remaining for effectively predicting the latent construct. Figure 5 illustrates the results for the model where the error variance was set at 59.

Model 3. Model 3 again uses the input model shown in Figure 5. In this model, however, the error variance of NR was fixed to 6, or 10.6% of the total variance of NR. All other parameters were left freed to be estimated. Changing the error variance of this one variable has a dramatic impact on the squared multiple correlations of nearly all of the other measured variables, as shown in Table 3. This is due largely in part because the number recognition variable of the Holzinger data has little to do with the

vocabulary and reading tasks which make up the other measured variables included in the analysis. By setting the error variance of NR to a low number, the entire analysis becomes more indicative of number recognition than vocabulary and reading, greatly reducing those variables' contributions to the analysis.

Model 4. Model 4 uses the same input model shown in Figure 5, with the error variance for GI set to 100, or 74.1% of the total item variance, and the error variance for PC set to 10, or 88.4% of the total item variance, with all other parameters left freed to be estimated. The changes in fit indices and squared multiple correlation can be seen in Table 3. While the effect on the squared multiple correlations here is not nearly as dramatic as in model 3, the measured variables attached to Var2 react differently. The squared multiple correlation of WM increases when compared to model 1, while the WC variable decreases.

Discussion

An examination of the results presented in Table 3 shows that changes in error estimates can greatly effect the squared multiple correlations (the proportions of non-error to total variance) of other variables included in the

model. The estimated squared multiple correlations for each of the variables in model 1 are less than the reliabilities for each of the variables. This is due to the fact that score reliability acts as an upper-limit to effect sizes (Capraro, Capraro, & Henson, 2001; Henson, 2001). Because the squared multiple correlations in SEM procedures are effect sizes, it is important that these squared correlations not exceed the reliability coefficients for those variables. If these squared multiple correlations did exceed the reliability coefficients, the use of SEM procedures make it possible to set the error variance equal to the un-reliable portion of the variable, forcing the effect size to remain equal to less than the reliability coefficient. This fact differentiates SEM from other GLM procedures; by taking into account measurement error, SEM allows the user to ensure that the effect sizes do not exceed the limits of reliability. In other GLM procedures, however, measurement error is not taken into account as part of the analysis. In fact, independent variables are considered to be essentially free of measurement error! This fact is by no means intended to deny the utility of non-SEM GLM procedures, rather it is meant to underscore the powerful

impact that measurement error can have on statistical results.

Because each of the estimated squared multiple correlations in model 1 were less than the variance reliabilities, setting the error variances equal to the unreliable portion of the measured variables (as done in model 2), serves to increase the individual effect sizes; however, an examination of the fit indices reveals that this is not a better solution to the original model. When all parameters are freed to be estimated, as in model 1, the results will be the best fit possible of the data to the model. Fixing any parameters to a value other than that models best estimate, given the sample, will always result in poorer fit indices.

In the case of model 3, a change in one of the error variances drastically changed the estimates of all other variables, changing the meaning of the analysis. While the changes in error in model 3 affected all variables in the same direction, the changes in error variance in model 4 had a different impact on the two variables contributing to Var2. Changes in error terms can effect the analysis in a variety of ways, at times increasing the effect sizes of some variables, at times decreasing the effect sizes of some variables, and at times doing both.

References

Arnold, M. E. (1996). Influences on and limitations of classical test theory reliability estimates. *Research in the Schools, 3*(2), 61-74.

Bacon, D. R., & others. (1995). Composite reliability in structural equations modeling. *Educational and Psychological Measurement, 55*, 394-406.

Bagozzi, R. P., Fornell, C., & Larcker, D. F. (1981). Canonical correlation analysis as a special case of a structural relations model. *Multivariate Behavioral Research, 16*, 437-454.

Bandalos, D. L. (1997). Assessing sources of error in structural equation models: The effects of sample size, reliability, and model misspecification. *Structural Equation Modeling, 4*, 177-192.

Capraro, M. M., Capraro, R. M., & Henson, R. K. (2001). Measurement error of scores on the mathematics anxiety rating scale across studies. *Educational and Psychological Measurement, 61*, 373-386.

Capraro, R. M., & Capraro, M. M. (2001). Commonality analysis: Understanding variance contributions to overall canonical correlation effects of attitude

toward mathematics on geometry achievement. *Multiple Linear Regression Viewpoints, 27*, 16-23.

Cohen, J. (1968). Multiple regression as a general data-analytic system. *Psychological Bulletin, 70*, 426-443.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrica, 16*, 197-334.

Dawson, T. E. (1998, April). *Structural equation modeling versus ordinary least squares canonical analysis: Some heuristic comparisons*. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA. (ERIC Reproduction Service No. ED 418 126)

Frederick, B. N. (1999). Partitioning variance in the multivariate case: A step-by-step guide to canonical commonality analysis. In B. Thompson (Ed.), *Advances in social sciences methodology* (Vol. 5, pp. 305-318). Stanford, CT: JAI Press.

Gorsuch, R. L. (1983). Factor analysis (2^{nd} ed.). Hillsdale, NJ: Erlbaum.

Gronlund, N. E., & Linn, R. L. (1990). *Measurement and evaluation in teaching* (6th ed.). New York: Macmillan.

Hancock, G. R. (1997). Correlation/validity coefficients disattenuated for score reliability: A structural

equation modeling approach. *Educational and Psychological Measurement, 57*, 598-606.

Hancock, G. R., & Nevitt, J. (1999). Bootstrapping and the identification of exogenous latent variables within structural equation models. *Structural Equation Modeling, 4*, 177-192.

Henson, R. K. (2001). Understanding internal consistency reliability estimates: A conceptual primer on coefficient alpha. *Measurement and Evaluation in Counseling and Development, 34*, 177-189.

Holzinger, K., & Swineford, F. (1939). *A study in factor analysis: The stability of a bi-factor solution* (No. 48). Chicago: University of Chicago. (Data on pp. 81-991)

Jöreskog, K. G. A. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika, 34*, 183-202.

Jöreskog, K. G. A. (1970). A general method for estimating of covariance structures. *Biometrika, 57*, 239-251.

Jöreskog, K. G. A. (1973). A general method for estimating a linear structural equation system. In A. S. Goldberger & O. D. Duncan (Eds.), *Structural equation models in the social sciences*, (pp. 85-112). New York: Seminar Press.

Jöreskog, K. G. A. (1977). Structural equation models in the social sciences: Specification, estimation, and testing. In P. R. Krishnaiah (Ed.). *Application of statistics,* (pp. 265-287). New York: North Holland.

Jöreskog, K. G. A., & Goldberger, A. S. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association, 10*, 631-639.

Jöreskog, K. G. A., & Sorbom, D. (eds.). (1978). *LISREL IV: Analysis of linear structural relationships by the method of maximum likelihood*. Chicago: National Education Resources.

Knapp, T. R. (1978). Canonical correlation analysis: A general parametric significance testing system. *Psychological Bulletin, 85*, 410-416.

Lomax, R. (1989). Covariance structure analysis: Extensions and developments. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 1, pp. 171-204). Greenwich, CT: JAI Press.

McArdle, J. J. (1984). On the madness in his madness: R. B. Cattell's contributions to structural equation modeling. *Multivariate Behavioral Research, 19*, 245-267.

Raines-Eudy, R. (2000). Using structural equation modeling to test for differential reliability and validity: An empirical demonstration. *Structural Equation Modeling, 7*, 124-141.

Raykov, T. (1998). A method for obtaining standard errors and confidence intervals of composite reliability for cogeneric items. *Applied Psychological Measurement, 22*, 369-374.

Reinhardt, B. (1996). Factors affecting coefficient alpha: A mini Monte Carlo study. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 4, pp. 3-20). Stamford, CT: JAI Press.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Thompson, B. (1991). A primer on the logic and use of canonical correlation analysis. *Measurement and Evaluation in Counseling and Development, 24*(2), 80-95.

Thompson, B. (1994). Guidelines for authors. *Educational and Psychological Measurement, 54*, 837-847.

Thompson, B. (2000). Ten commandments of structural equation modeling. In L. Grimm & P. Yarnold (Eds.), *Reading and understanding more multivariate statistics*

(pp. 261-284). Washington, DC: American Psychological Association.

Wilkinson, L., & APA Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54*, 594-604. [reprint available through the APA Home Page: http://www.apa.org/journals/amp/amp548594.html]

Table 1
Summary of Regression Analysis for Variables Predicting Word Meaning (N=145)

	Unstandardized Coefficients				
	B	Std. Error	Beta	t	Sig.
Constant	-7.262	4.853		-1.496	.137
GI	.306	.044	.449	6.891	.000
PC	1.020	.155	.433	6.597	.000
NR	.007	.056	.007	.131	.896

Note. R Square=.635

Table 2

Correlation Matrix of t5, t6, t15

Variables	T5	T6	T15
T5	1.000	*.622	*.219
T6		1.000	*.249

Note. N=145.
* Correlation is significant at the 0.01 level.

Table 3

SEM Model Squared Multiple Correlations and Fit Indices

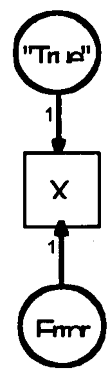
	Squared Multiple Correlation					Chi-			
Model	GI	PC	NR	WM	WC	Square	df	RMSEA	AGFI
1	.67	.64	.07	.77	.42	5.1	5	.011	.960
2	**.79**	**.64**	**.13**	**.86**	**.50**	79.5	10	.220	.772
3	.09	.10	**.89**	.09	.06	259.9	6	.542	-.154
4	**.43**	**.38**	.06	.92	.34	62.2	7	.234	.685

Note. Values in **bold** indicate variables whose error terms have been fixed.

Table 4

Variable Reliability Coefficients and Variances

Coefficient		Variance		
	Alpha	Total	Non-Error	Error
GI	0.8077	134.97	109.0153	25.95473
PC	0.6507	11.315	7.362671	3.95233
WC	0.5802	28.505	16.5386	11.9664
WM	0.8701	62.727	54.57876	8.148237
NR	0.5070	56.496	28.64347	27.85253



<u>Figure 1.</u> Partitioning of measured variable variance into error and non-error expressed in structural equation modeling.

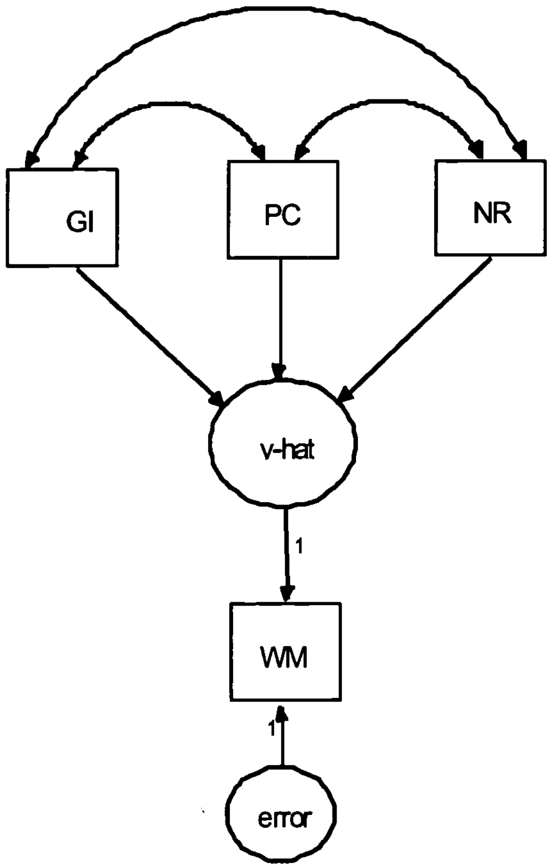


Figure 2. Regression analysis modeled in SEM.

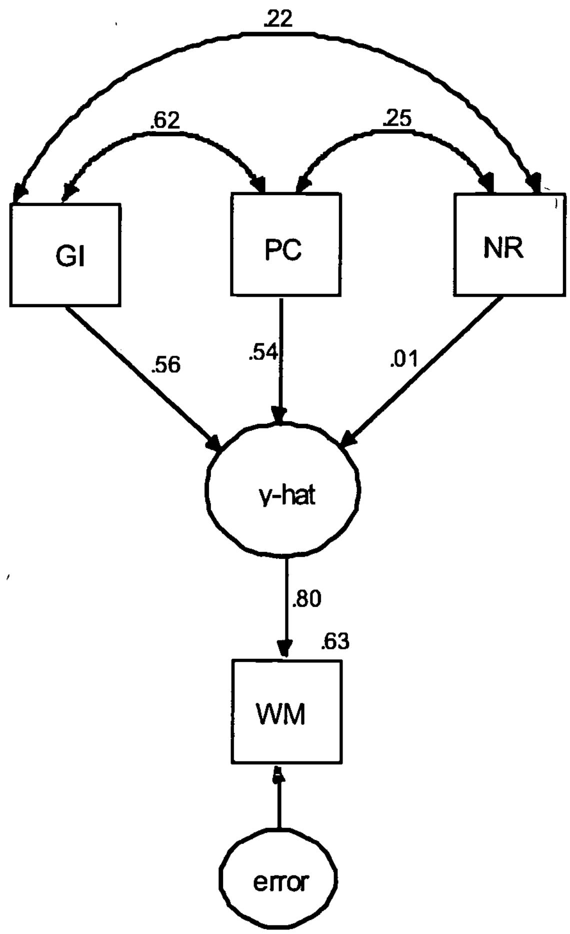


Figure 3. Standardized output from regression analysis.

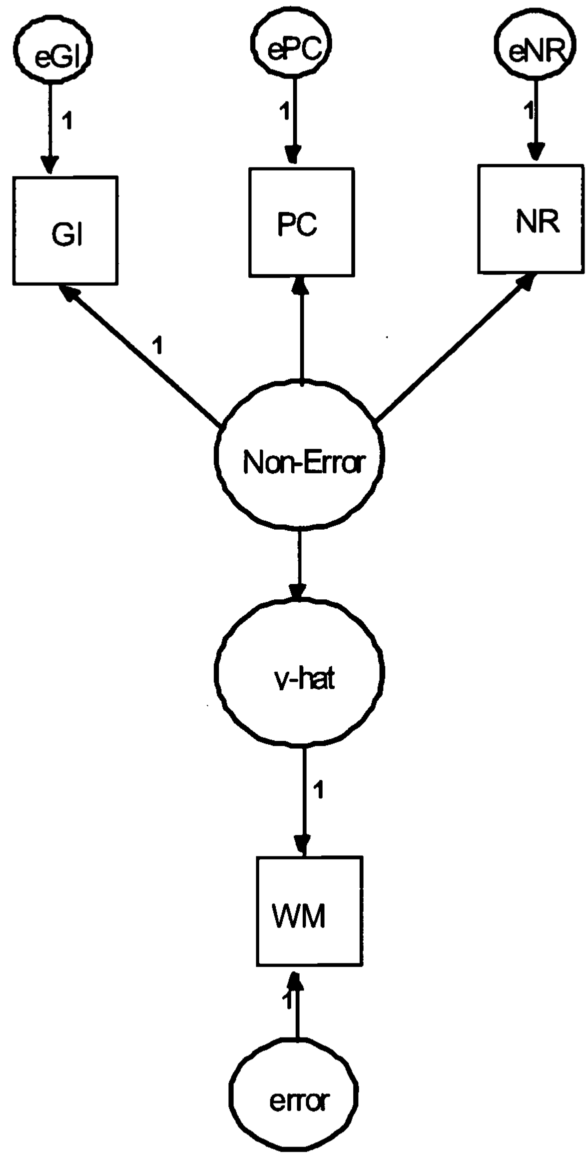


Figure 4. Regression model with the independent variables partitioned into error and non-error.

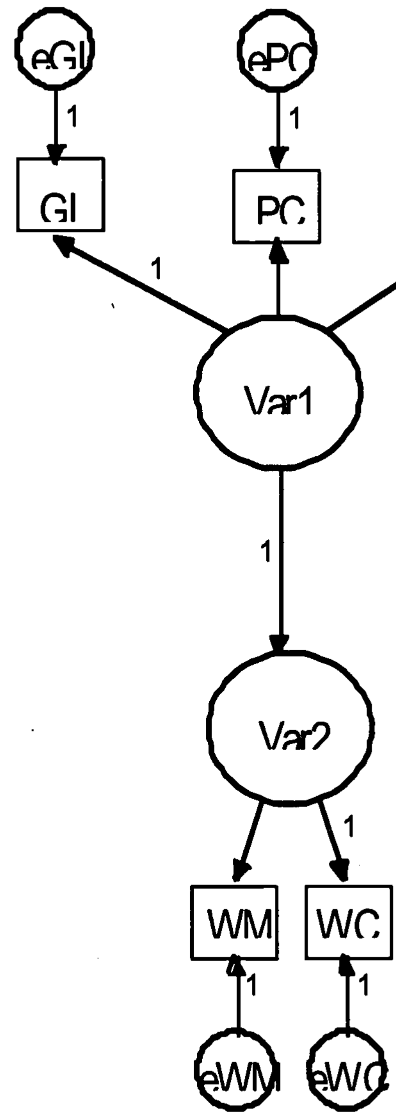


Figure 5. Input model for SEM heuristic examples.

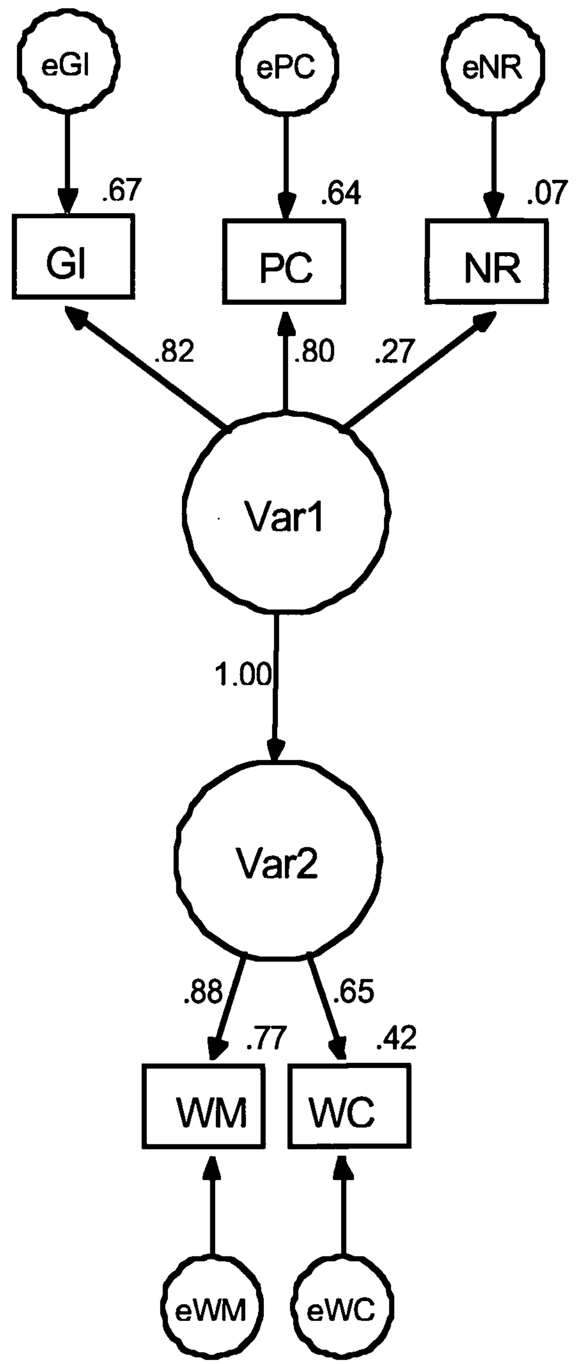


Figure 6. AMOS output with all parameters freed to be estimated.

Appendix A

SPSS Syntax

```
Comment creates the variable school and assigns either a 1 or 2.
compute school=1.
if (ID gt 200) school=2.
execute.

Comment computes a regression analysis for the data for comparison to
the output from AMOS.
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT wm
  /METHOD=ENTER gi pc nr  .
REGRESSION
  /SELECT= school EQ 2
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT t9
  /METHOD=ENTER t5 t6 t15  .

Comment selects only those cases with School=2.
USE ALL.
COMPUTE filter_$=(school=2).
VARIABLE LABEL filter_$ 'school=2 (FILTER)'.
VALUE LABELS filter_$  0 'Not Selected' 1 'Selected'.
FORMAT filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE .

Comment Computes the correlation matrix for comparison to the Reg.
Analysis.
CORRELATIONS
  /VARIABLES=t5 t6 t15
  /PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE .
```

Appendix B

Variable Data from Holzinger & Swineford (1939)

GI	*PC*	*WC*	*WM*	*NR*
46	10	22	10	91
43	8	30	10	81
36	11	27	19	84
38	9	25	11	84
51	8	28	24	98
42	10	28	18	86
69	17	42	41	95
35	10	29	11	97
32	11	35	8	84
39	9	27	16	74
27	10	25	13	95
27	1	29	11	100
29	10	30	14	90
35	5	11	10	94
56	14	29	26	96
37	7	18	11	97
48	11	29	18	86
65	10	23	35	88
49	8	26	20	90
59	13	33	36	98
56	5	27	11	84
56	13	39	25	95
50	14	28	24	87
25	7	17	4	94
29	8	29	13	84
39	9	26	17	79
55	13	26	19	79
24	9	24	7	92
44	4	30	16	89
49	9	26	17	91
46	9	23	13	77
51	9	30	21	90
33	9	27	20	91
18	5	21	4	89
31	8	30	8	74
41	9	31	15	96
48	9	28	20	88
20	2	10	2	85

42	6	35	13	81
52	19	36	33	101
63	15	32	25	90
47	9	24	24	76
29	9	24	8	80
37	8	29	10	87
41	14	24	11	92
47	11	27	23	90
20	6	15	10	73
44	11	32	18	100
28	8	26	17	84
53	10	32	18	87
33	7	21	12	88
42	7	27	13	93
32	5	24	5	95
50	6	34	13	83
54	14	31	20	88
36	7	27	14	98
45	9	28	17	85
56	18	31	27	91
30	8	25	6	89
40	10	24	15	86
56	18	37	31	84
23	7	22	7	95
48	9	36	16	88
16	5	12	4	94
29	6	25	6	73
39	13	22	15	92
32	10	30	16	84
43	15	26	19	84
41	9	28	14	83
49	8	29	12	92
46	9	24	10	96
39	10	26	17	86
40	8	25	17	78
61	15	36	39	96
31	7	19	19	94
34	9	27	22	96
33	15	32	30	95
35	8	26	13	81
43	5	28	11	82
41	9	28	11	95
50	13	29	26	79
39	16	28	15	92
41	15	31	21	86
55	9	29	23	84
55	13	27	17	110

54	10	34	27	89
48	11	36	18	92
58	11	37	12	104
44	8	33	15	89
61	10	33	19	82
48	10	22	19	98
60	12	30	23	83
50	11	31	19	85
54	9	24	12	90
51	12	32	15	94
84	18	43	38	112
60	15	29	30	86
64	12	23	33	86
72	16	28	32	104
36	8	33	19	86
43	7	23	19	84
78	17	35	30	101
45	10	26	16	94
55	10	29	19	88
50	10	37	23	88
38	6	27	9	88
51	12	33	17	103
57	13	31	26	79
55	14	28	29	92
51	12	33	27	96
59	18	34	33	97
60	14	29	24	79
66	14	36	29	96
41	10	25	13	95
37	14	33	14	100
48	10	30	17	95
46	8	33	15	89
53	8	29	13	86
44	7	34	10	94
40	4	25	17	87
52	9	29	14	86
37	12	21	15	90
42	5	23	16	87
39	8	24	13	111
36	7	28	8	74
58	15	34	23	85
44	8	22	12	91
42	12	32	19	79
42	8	30	12	92
55	10	26	10	85
48	13	33	32	89
43	7	25	11	76

42	9	31	23	99
53	10	26	10	99
37	8	28	6	92
31	4	27	11	85
44	10	25	14	82
51	10	28	14	99
45	9	26	11	96
49	9	31	20	92
49	8	27	7	75
31	7	23	7	87
55	11	32	30	92
48	11	33	14	87
51	11	39	22	90

TM033706

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC

Reproduction Release

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: A Tutorial on How relaibility is Estimated Within Structural Equation Models	
Author(s): Robert M. Capraro & James M. Graham	
Corporate Source:	Publication Date: 2/14/02

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, Resources in Education (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign in the indicated space following.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY SAMPLE TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY SAMPLE TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY SAMPLE TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
Level 1	Level 2A	Level 2B
☑	☐	☐
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g. electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche, or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Signature:	Printed Name/Position/Title: Robert M.Capraro/Professor/Mathematics Education	
Organization/Address: Texas A&M Univeristy 4232 TAMU College Station, TX 77843-4232	Telephone: 979-845-8007	Fax: 979-845-9663
	E-mail Address: rcapraro@coe.tamu.edu	Date: 2/14/2002

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:	
ERIC Clearinghouse on Assessment and Evaluation **1129 Shriver Laboratory (Bldg 075)** **College Park, Maryland 20742**	**Telephone: 301-405-7449** **Toll Free: 800-464-3742** **Fax: 301-405-8134** **ericae@ericae.net** **http://ericae.net**

EFF-088 (Rev. 9/97)